DATA FUSION COMMENTS – George Huffman (1/2)

Note that NOAA is already involved in the PMM merged data products

- critical input precipitation datasets
- part of the team developing Integrated Multi-satellitE Retrievals for GPM (IMERG)
  - code-level integration based on KF-CMORPH, PERSIANN-CCS, TMPA

Error estimates are key to intelligent fusion

- precipitation is statistically gnarly because of intermittency
- each sensor has particular error characteristics
  - <insert massive laundry lists for satellite, radar, gauge, and model!!>
  - relative utility of each varies widely by surface type and weather regime
- we have to get much better at compact, useful uncertainty statements
  - for the science
  - for diverse users

Ensemble estimation may be useful

- note examples of ensemble NWP and flood modeling
- sidesteps issue of how to combine errors
- at a minimum, having several different algorithms can be helpful

DATA FUSION COMMENTS – George Huffman (2/2)

An additional error issue is that (some) users need the data at finer scales than the observations

- statistical downscaling in not yet satisfactory

We have to figure out using models

- in principle they are better in "difficult" conditions
  - complex terrain, winter, fine scales
- note Australian Short Term Ensemble Prediction System (STEPS)
  - fuse fine-scale radar and model nowcasts

Don't forget the basic inputs

- satellite precip estimates over land, coast still lag estimates over ocean
- all sensors struggle in winter, polar conditions
- the new best practice is episodic reprocessing for entire datasets
- sensor replenishment is vital
  - gauges, particularly snow-capable
  - satellite microwave sensors are sparse, uncoordinated
  - microwave imagers are losing launch opportunities

IMERG DESIGN – Notional Requirements

Resolution – 0.05° ~0.1° [i.e., roughly the resolution of microwave, IR footprints]

Time interval – 30 min. [i.e., the geo-satellite interval, then aggregated to 3 hr]

Spatial domain – global, initially covering 60° N-60° S

Time domain – 1998-present; later explore entire SSM/I era (1987-present)

Product sequence – early sat. (~4 hr), late sat. (~12 hr), final sat.-gauge (~2 months after month) [more data in longer-latency products] unique in the field

Instantaneous vs. accumulated – accumulation for monthly; instantaneous for half-hour

Sensor precipitation products intercalibrated to TRMM before launch, later to GPM

Global, monthly gauge analyses including retrospective product – explore use in submonthly-to-daily and near-real-time products; unique in the field

Error estimates – still open for definition; nearly unique in the field

Precip type – simple probability of liquid/mixed/solid; nearly unique in the field

Embedded metadata fields showing how the estimates were computed

Operationally feasible, robust to data drop-outs and (strongly) changing constellation

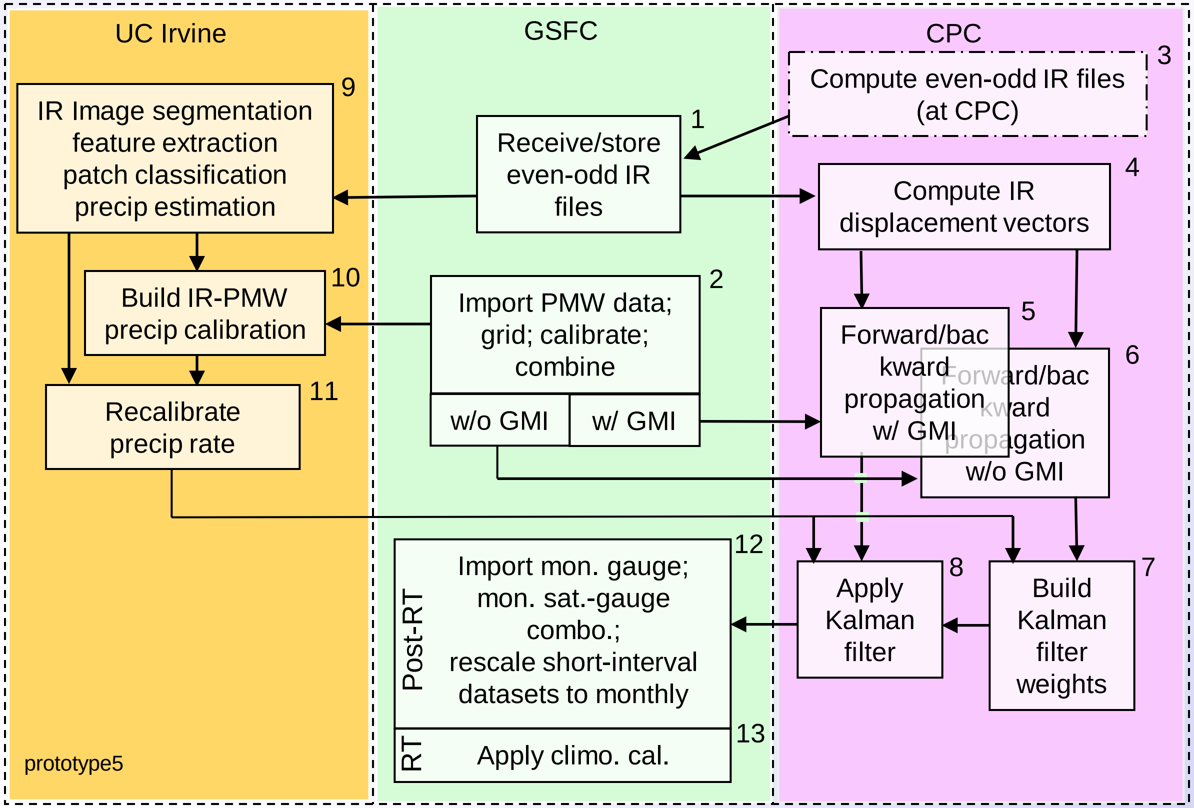
Output in HDF5 v1.8 – compatible with NetCDF4

Archiving and reprocessing for near- and post-RT products; nearly unique in the field

IMERG DESIGN – Box Diagram

The flow chart shown is for the final product

- institutions are shown for module origins, but
- package will be an integrated system
- "the devil is in the details"
- (near-)RT products will use a cut-down of this processing
- we will examine the utility of climatological calibrations for the (near-)RT products



IMERG DESIGN – Data Fields

Output dataset needs to include intermediate data fields

- users and developers require
  - traceability of processing, and
  - support for algorithm studies

0.1° global CED grid

- 3600x1800 = 6.2M boxes
- Fields are 1-byte integer or or scaled 2-byte integer / 4-byte real

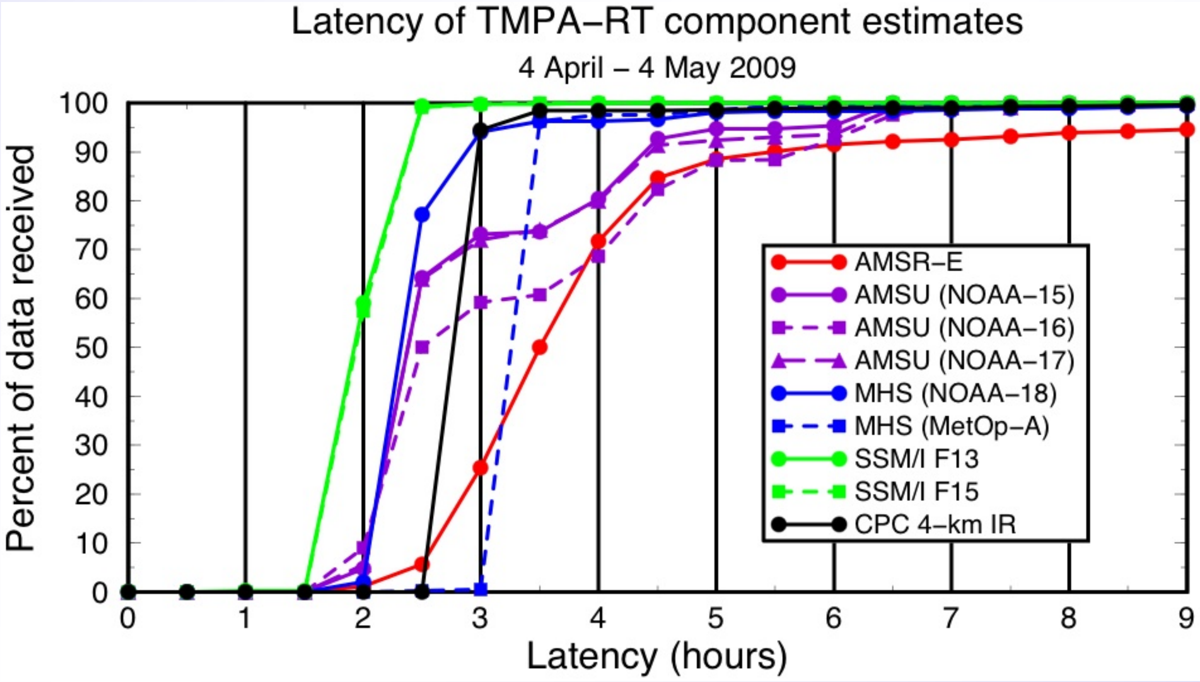
"User" fields in italics, darker shading

Half-hourly data file (Early, Late, Final)	
1	*Instantaneous precipitation - calibrated*
2	Instantaneous precipitation - uncalibrated
3	*Precipitation error*
4	PMW precipitation
5	PMW source 1 identifier
6	PMW source 1 time
7	PMW source 2 identifier
8	PMW source 2 time
9	IR precipitation
10	IR KF weight
11	*Precipitation type*
Monthly data file (Final)	
1	*Sat-Gauge precipitation*
2	*Sat-Gauge precipitation error*
3	Gauge relative weighting
4	*Precipitation type*

IMERGE DESIGN – Multiple Runs

Multiple runs serve different users' needs for timeliness

- more delay usually yields a better product
- pioneered in TMPA



Final – after the best data are assembled; research users

- driver is precip gauge analysis
- GPCC gauge analysis is ~2 months after the month

Late – wait for full multi-satellite; crop, flood, drought analysts

- driver is waiting for microwave data for backward propagation
- expect delay of 12-18 hr

Early – get a first approximation; flood, short-range forecasting users

- current input data latencies support ~4-hr delay
- can't support truly operational users (flash flood, nowcasting), who need < 3 hr

PRECIP CONSTELLATION

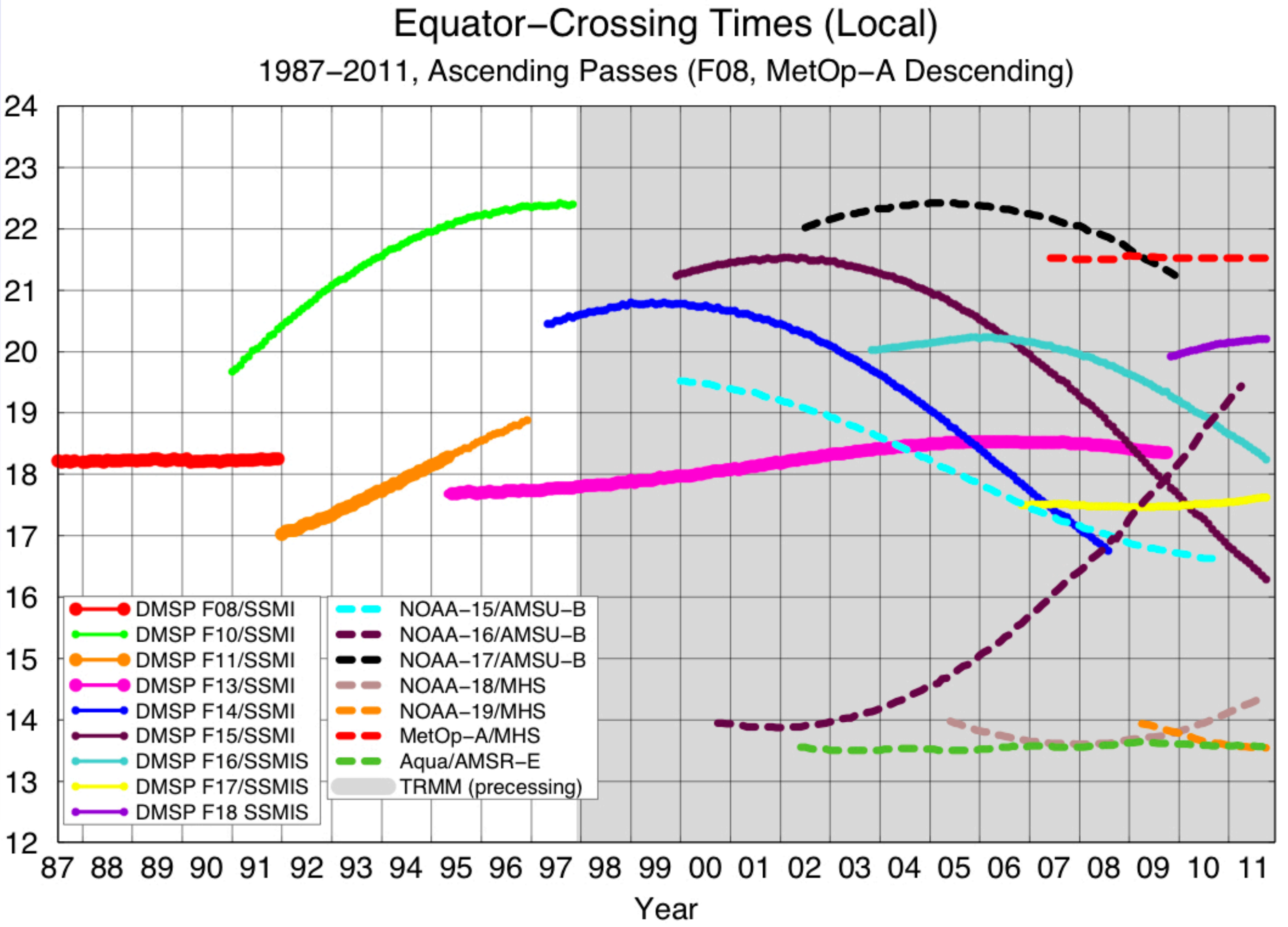
Passive microwave are divided into imagers and sounders

Imagers are better for precip because they have

- finer resolution
- consistent footprint size
- polarized channels

Coverage is sparse and depends on legacy satellites, even for sounders

SSMIS brings imager benefits to sounder channels



Thickest lines denote GPCP calibrator.

Image by Eric Nelkin (SSAI), 27 October 2011, NASA/Goddard Space Flight Center, Greenbelt, MD.